The inception cohort comprised children aged 4-6 years who were either in treatment or presenting for assessment at six different centers that serve preschool children with developmental disabilities in Southern Ontario. Details on ascertainment are available from previous publications (Bennett et al., 2008). Written informed consent was obtained from the parents of each child and the children’s assent was obtained by judging their willingness to participate in the assessment. Approval for the study was obtained from McMaster University’s research ethics board.

Children aged 4-6 who were included in this cohort of high-functioning children with ASD, which started in 1987 demonstrated at least one example of a DSM-III-R criterion of impairment in each domain of social reciprocity and verbal and nonverbal communication, as well as at least one example of a repetitive, restricted or stereotyped behavior. Participants also needed to demonstrate a Leiter IQ score above 68 or a Stanford-Binet score above 70 (i.e., above the cut-off for global developmental delay). Children were excluded if they were untestable or if they did not meet the cognitive cut-off criteria.

The enrolment cohort included 68 high-functioning preschool children with ASD, all with IQs above 68 on the Leiter or 70 on the Stanford-Binet scales. The families were contacted approximately two years later and the Test of Language Development-2 (TOLD-2; Newcomer & Hammill, 1988) was administered. Times 1, 2 and 3, as described in the current study, were conducted approximately two, six and eight years respectively, after the enrolment phase. To ensure blindness, a different research assistant than the one who did the original assessment conducted each follow-up assessment.

The “Reading the Mind in the Eyes Test” was included in the battery of tests at Time 2 (six years after initial interview and the enrolment phase, approximate ages 10-12 years) as a measure of ToM (Baron-Cohen, Jolliffe, Mortimore, & Robertson, 1997). At the time of inclusion of this measure, 29 children had already completed their age 10-12 year assessment visit so this measurement tool could not be administered for that time point. The cohort for the current study therefore comprised 39 high-functioning children with ASD who were able to reach a basal score on the “Eyes Test”. Analyses were performed to compare this subgroup of 39 children to the 29 children who had been excluded with respect to key demographic variables and characteristics to determine whether, this subsample remained representative of the larger group. There were no differences found on t-test and chi-squared analyses of gender, age at entry into the study, diagnostic group (high-functioning autism versus Asperger Syndrome), performance IQ, TOLD-2 or Vineland scores at Time 3. The outcome measure of adaptive functioning was taken roughly two years later (at Time 3) when the children were between 14-16 years of age. Only the socialization and communication subscale scores are used in this analysis since it was hypothesized that these would be most strongly associated with ToM abilities.
Measures

Autism Diagnostic Interview (ADI): The ADI is a semi-structured interview designed to make a diagnosis of autism according to both the DSM-III-R and ICD-10 criteria and was administered during the enrolment phase. The interview has good reliability (intra-class correlation for multiple raters ranges between .94 to .97) (Lord, Rutter, & LeCouteur, 1994).

Time 1 measures (age 6-8 years)

Arthur Adaptation of the Leiter Performance Scales, (Levine, 1986): This is a standardized measure of nonverbal problem solving (mean = 100, s.d. = 15). The Leiter is widely used with ASD and other language-impaired children. The Leiter is especially appropriate to the population under study because it does not require verbal instructions for administration and is highly correlated with WISC-R IQ.

The Test of Language Development-2 (TOLD-2; Newcomer & Hammill, 1988): The Grammatic Completion, Grammatic Understanding and Vocabulary subtests of the TOLD-2 were used to measure grammatical comprehension and usage. Standard scores (mean = 10, SD=3) were calculated for each child. If the child was mute or unable to obtain a basal score on the TOLD s/he was given an imputed score of one below the lowest score of a child who was able to complete the test. This was done to minimize the number of missing values, and to provide a reasonably valid, if conservative, estimate of abilities. Seven of this study sample of 39 children received imputed scores.

Time 2 measures (age 10-12 years)

The “Reading the Mind in the Eyes Test” (“Eyes test”; Baron-Cohen et al., 1997): This task requires an individual to view 30 copies of magazine photographs of male and female faces, truncated to include only the eye regions and to choose from different terms the one that best depicts the mental or emotional state that is expressed. The Eyes test
has been shown to distinguish between high-functioning individuals with ASD and typically developing controls (Baron-Cohen et al., 1997). A revised version was later developed (Baron-Cohen, Wheelwright, & Hill, 2001). In the current study, the number of options for mental- or emotional-state terms was reduced from 4 to 2 based on pilot work that suggested the 4-choice option was too difficult for this sample. The two possible responses comprised the correct answer and a second chosen at random from the remaining three options.

**Time 3 measures (age 12-14 years)**

**Vineland Adaptive Behavior Scales, 1st edition, (VABS; Sparrow, Balla, & Cicchetti, 1984):** The Vineland is a semi-structured interview designed to assess personal and social sufficiency of individuals from birth to adulthood. Specifically, it measures strengths and weaknesses in socialization, communication, and activities of daily living in children aged 0 to 18 years of age (mean = 100, s.d. = 15). The Vineland requires an interview with a respondent familiar with the individual’s behaviour.

**References**


